

10/758,486

IBM/265

In the Specification

Under the Brief Description of Drawings, page 5, after line 2, please add the following:

Fig. 5 illustrates a linked list structure for storing frequent characters.

Fig. 6 illustrates a process for identifying the number of results of a search string.

10/758,486

IBM/265

The two paragraphs beginning on page 11, line 12 to page 12, line 11, are replaced with the following:

A linked list structure such as shown in Fig. 5[[4]] has the advantage that the number of frequent characters and frequent subsequent characters may be flexibly changed. For example, a linked list structure such as in Fig. 5[[4]] could store those characters that occur more than x% of the time in each character position, and for each frequent character, those subsequent characters that occur more than y% of the time; the linked lists would then be larger or smaller as needed to capture the most frequent characters based upon the desired threshold percentages x and y.

Using this structure, statistics for the likely number of rows matching a pattern matching predicate can be readily generated. For example, referring to Fig. 6[[5]], the process for identifying the number of results of a search string starting at position n, can be described. In a first step 300, the first potential position of a search string in the key values of the relevant attribute, is identified. Then, a sequence of steps 302 through 344 is performed to estimate the count of results matching the search string at that position. After these steps, in step 346 it is determine whether there is another potential position of the search string in the key values, and if so, in step 348, that potential position is selected and the steps 302 through 344 are repeated to estimate the count of results matching the search string at that position. After all potential positions of the search string in the key values have been evaluated, the search process is done, and in step 350 the counts at each potential position are accumulated to produce the estimate of the total count of results for the search string.